Torin Flanagan 1169130

TDF008@student.usc.edu.au

Abstract

The goal of this report was to develop a comprehensive data management plan for Itikan, addressing all required areas including data collection, metadata, ethics and compliance, storage and backup, preservation, data sharing, and data governance. Each section of the plan begins with a general discussion or definition of the topic, followed by a specific application to Itikan’s needs. The result is a detailed plan that covers all essential aspects of managing the company’s e-commerce data effectively.

Data Management Plan

Itikan and E-Commerce Data

Contents

[1. Introduction 3](#_Toc180599877)

[1.1 What is a Data Management Plan? 3](#_Toc180599878)

[1.2 Why use a Data Management Plan? 3](#_Toc180599879)

[1.3 Data Management Plan Framework 3](#_Toc180599880)

[2. Data Collection 4](#_Toc180599881)

[2.1 What is Data Collection? 4](#_Toc180599882)

[2.2 Why is Data Collection Important? 4](#_Toc180599883)

[2.3 Processes Used for the Collection of Data 4](#_Toc180599884)

[2.4 What Data Sources Are Used in a Data Management Plan? 5](#_Toc180599885)

[2.5 Decision for Itikan E-Commerce Company 5](#_Toc180599886)

[3. Metadata 6](#_Toc180599887)

[3.1 What is Metadata? 6](#_Toc180599888)

[3.2 Why is Metadata Important? 6](#_Toc180599889)

[3.3 Metadata Management 6](#_Toc180599890)

[3.4 Standards for Metadata in Australia 7](#_Toc180599891)

[3.5 Decision for Itikan E-Commerce Company 7](#_Toc180599892)

[4. Master Data 8](#_Toc180599893)

[4.1 What is Master Data? 8](#_Toc180599894)

[4.2 Why is Master Data Important for a Company? 8](#_Toc180599895)

[4.3 Decision for Itikan E-Commerce Company 8](#_Toc180599896)

[5. Data Ethics 9](#_Toc180599897)

[5.1 What are Data Ethics? 9](#_Toc180599898)

[5.2 Why are Data Ethics Important? 9](#_Toc180599899)

[5.3 Decision for Itikan E-Commerce Company 9](#_Toc180599900)

[6. Legal Compliance 10](#_Toc180599901)

[6.1 What is Legal Compliance 10](#_Toc180599902)

[6.2 Regulatory Frameworks, Data Protection Policies, and Audits 10](#_Toc180599903)

[6.3 Decision for Itikan E-Commerce Company 10](#_Toc180599904)

[7. Storage and Backup 11](#_Toc180599905)

[7.1 Why is Storage and Backup Necessary? 11](#_Toc180599906)

[7.2 Options for Storage and Backup 11](#_Toc180599907)

[7.3 Decision for Itikan E-Commerce Company 12](#_Toc180599908)

[8. Preservation 12](#_Toc180599909)

[8.1 What is Data Preservation? 12](#_Toc180599910)

[8.2 How Long Should Data be Stored For? 12](#_Toc180599911)

[8.3 Decision for Itikan E-Commerce Company 13](#_Toc180599912)

[9. Data Sharing 13](#_Toc180599913)

[9.1 What is Data Sharing? 13](#_Toc180599914)

[9.2 Internal Data Sharing 14](#_Toc180599915)

[9.3 External Data Sharing 14](#_Toc180599916)

[9.4 Levels of Authority 14](#_Toc180599917)

[9.5 Decision for Itikan E-Commerce Company 14](#_Toc180599918)

[10. Data Governance 15](#_Toc180599919)

[10.1 What is Data Governance? 15](#_Toc180599920)

[10.2 Data Policies 16](#_Toc180599921)

[10.3 Data Standards 16](#_Toc180599922)

[10.4 Decision for Itikan E-Commerce Company 16](#_Toc180599923)

[11. Data Security 17](#_Toc180599924)

[11.1 What is Data Security? 17](#_Toc180599925)

[11.2 Data Security Strategies 18](#_Toc180599926)

[11.3 Decision for Itikan E-Commerce Company 18](#_Toc180599927)

[12. Data Disposal 18](#_Toc180599928)

[12.1 Why is Data Disposal Important? 18](#_Toc180599929)

[12.2 How is Data Organised for Disposal? 19](#_Toc180599930)

[12.3 Decision for Itikan E-Commerce Company 19](#_Toc180599931)

[13. Data Integration 19](#_Toc180599932)

[13.1 How Does Data Integrate? 19](#_Toc180599933)

[13.2 Why is Data Integration Necessary? 20](#_Toc180599934)

[13.3 Decision for Itikan E-Commerce Company 20](#_Toc180599935)

[14. Monitoring and Reviewing of Data 20](#_Toc180599936)

[14.1 Why is it Important to Monitor and Review data 20](#_Toc180599937)

[14.2 Decision for Itikan E-Commerce Company 20](#_Toc180599938)

[15. Conclusion 20](#_Toc180599939)

[16. References 21](#_Toc180599940)

[17. Appendix 24](#_Toc180599941)

[Appendix 1 (Sample Marketing Data) 24](#_Toc180599942)

**I acknowledge the use of Chat GPT in this assessment for proof reading my work.**

# Introduction

The CEO of Itikan, impressed by a previous dashboard project for Jonathon Briggs, has requested the creation of a Data Management Plan (DMP). His marketing team works with external agencies on campaigns, but the company lacks a formal data management strategy. Currently, sensitive data like customer details and metrics are scattered across databases and Excel files, putting information at risk. A DMP will ensure appropriate handling of e-commerce data, safeguarding sensitive information and optimising its use for marketing.

## 1.1 What is a Data Management Plan?

A DMP is a document that outlines how data will be handled during and after a project (Michener, 2015). Key elements include data description, impact, format, privacy protection, access, preservation, and transfer of responsibility (Thoegersen, 2015).

## 1.2 Why use a Data Management Plan?

A DMP is essential for any data-driven project, as it defines data policies, access rules, management practices, and responsibilities (Australia Research Data Commons, 2022). It helps Itikan allocate the resources, tools, and expertise needed for managing its data (CESSDA, 2017).

## 1.3 Data Management Plan Framework

Diagram

Description automatically generated The framework consists of foundational practices (principles for data management) and advanced practices (technologies and guidelines for using data).

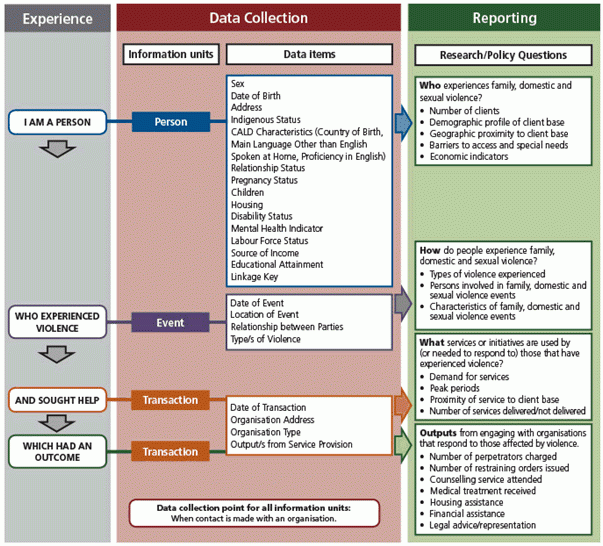
*Figure 1: Data Management Practices Hierarchy (Lang, 2024).*

# Data Collection

## 2.1 What is Data Collection?

Data collection is the first step in a DMP, outlining how data is collected, stored, and accessed (IBM, n.d.). It serves multiple purposes, such as improving business processes, understanding customer behaviour, measuring performance, and aiding decision-making. Data collection follows a strict ETL (Extract, Transform, Load) process, ensuring raw data is sourced, formatted consistently, and stored in a data warehouse or database.

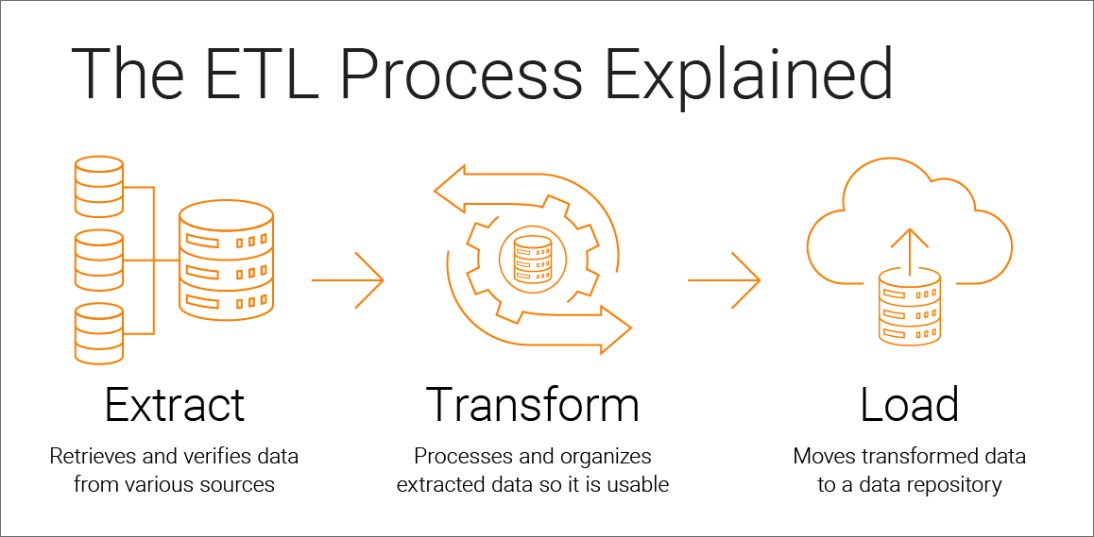
## 2.2 Why is Data Collection Important?

In today's technological era, data collection is crucial for decision-making, improving efficiency, and ensuring long-term success. Accurate data enables informed decisions in product development, marketing, resource allocation, and identifying new business opportunities (Baker, 2023).

*Figure 2: Data Collection and Reporting Framework (Australian Bureau of Statistics, 2014).*

## 2.3 Processes Used for the Collection of Data

The ETL (Extract, Transform, Load) process is a widely adopted method for data collection and management. In the extract phase, data is gathered from internal and external sources through methods like surveys and transactions. During transformation, the data is cleaned and standardised for consistency. Finally, in the load phase, the transformed data is stored in databases like SQL, NoSQL, or cloud-based services, offering better scalability and reliability than spreadsheets.

*Figure 3: Extract, Transform, and Load process for data collection (Informatica, 2024).*

## 2.4 What Data Sources Are Used in a Data Management Plan?

Data collection in a DMP gathers key information from internal and external sources. Internal sources provide primary data, while external sources offer secondary data. The table below outlines the types of data collected from within the business and the broader market:

|  |  |
| --- | --- |
| **Internal Sources (Business)** | **External Sources (Market)** |
| Operations | Surveys |
| Maintenance | Questionnaires |
| Personnel | Research |
| Finance | Customer Feedback |

*Table 1: Types of Data Collection for Different Data Sources.*

## 2.5 Decision for Itikan E-Commerce Company

Based on the discussions about data collection, we must now decide on Itikan’s approach for its DMP. Data collection is critical as Itikan’s DMP will rely heavily on it.   
From appendix 1, it shows existing data such as ‘Customer ID,’ ‘Customer Name,’ and transaction-related information. Additional headings like ‘Email,’ ‘Phone Number,’ and ‘Average Order Value’ could be added to further increase data collection. Meanwhile, data such as ‘Email’ and ‘Phone Number’ can be collected during customer purchases, while ‘Average Order Value’ can be calculated from dividing 'Amount 2 Years' by 'Frequency 2 Years'. It is important to note that all collected data will follow a consistent format.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Email** | **Phone Number** | **Frequency 1 Year** | **Amount 1 Year** | **Average Order Value** |
| JennyRobinson@gmail.com | 0455 956 876 | 4 | $101.54 | $26.24 |
| AnnFoster@outlook.com | 0406 876 321 | 6 | $256.76 | $40.17 |
| MarkClapham@hotmail.com | 0456 856 555 | 5 | $375.85 | $94.63 |

*Table 2: Example of extended data set heading values.*

# Metadata

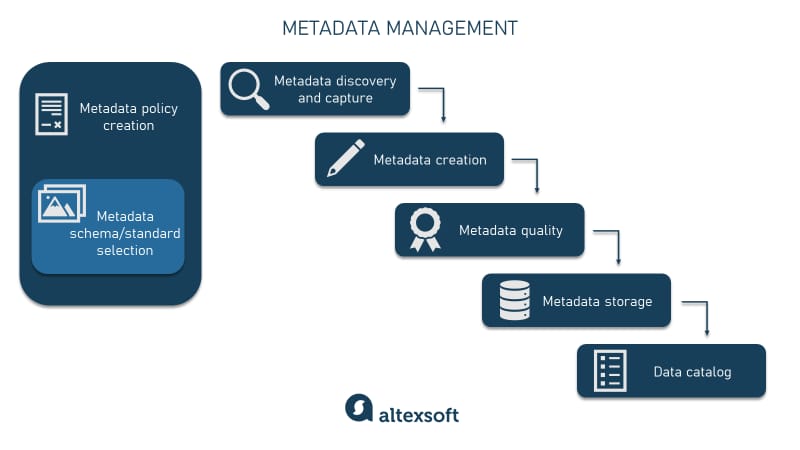
## 3.1 What is Metadata?

Metadata is data about data, providing higher-level information describing the content, context, quality, structure, and accessibility of a dataset (Mayernik, 2015).

## 3.2 Why is Metadata Important?

Metadata is essential in a DMP as it clarifies the dataset's content, making it easier to use, archive, preserve, and reuse. Well-defined metadata simplifies data management compared to datasets without clear definitions (Singh & Madalli, 2023).

## 3.3 Metadata Management

Metadata has become essential for data integration, change management, and analytics, requiring dedicated oversight by repository managers (Sen, 2004). As metadata management has evolved from implicit practices to explicit management, companies now implement comprehensive plans to ensure proper storage and handling.

*Figure 4: Steps involved for metadata management (Alexsoft, 2022).*

## 3.4 Standards for Metadata in Australia

The Australian Government Recordkeeping Metadata Standard (AGRkMS) was introduced in 1999 to establish a framework for standardising metadata. Ongoing efforts in ISO and Standards Australia resulted in three AS ISO 23081 specifications:

* **ISO 23081-1:2017**: Defines principles for metadata requirements (Cunningham, n.d.).
* **ISO 23081-2:2021**: Covers implementation, management, and conceptual models for metadata (Cunningham, n.d.).
* **ISO/TR 23081-3:2011**: Provides a self-assessment checklist to evaluate metadata schema compliance (Cunningham, n.d.).

## 3.5 Decision for Itikan E-Commerce Company

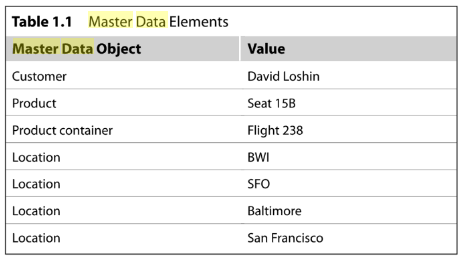
Metadata management at Itikan will be customised to the company’s processes for acquiring, storing, analysing, and using such customer data. As shown in Figure 4, the plan includes capturing, creating, ensuring quality, storing metadata, and generating a data dictionary. Itikan’s metadata must also comply with Australian standards, below is the data dictionary created for Itikan’s metadata:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Description** | **Data Type** | **Length** | **Range** | **Sensitive** |
| Customer ID | Unique identification for customer | Integer | 3 | 211-526 | Yes |
| Customer Name | Name associated with Customer ID | String | 40 | N.A. | Yes |
| Recent Purchase | When the customer last purchased an item from Itikan | Date | 7 | 1/1/2023-31/12/2023 | Yes |
| Amount | How much the customer spent on their most recent purchase | Currency | 5 | $0.00-$999.99 | No |
| Street Number | Customer’s address | Integer | 3 | 0-999 | Yes |
| Street | Customer’s address | String | 40 | N.A. | Yes |
| State | Customer’s address | String | 12 | N.A | Yes |
| Post Code | Customer’s address | Integer | 4 | 0000-9999 | Yes |
| First Purchase | When customer first purchased something from Itikan | Date | 7 | 1/1/2022-31/12/2024 | No |
| Frequency 6 Months | How frequent the customer purchases items within 6 months | Integer | 1 | 0-9 | Yes |
| Frequency 2 Years | How frequent the customer purchases items within 2 years | Integer | 1 | 0-9 | Yes |
| Amount 6 Months | How much the customer has spent at Itikan in the past 6 months | Currency | 6 | $0.00-$9999.99 | No |
| Amount 2 Years | How much the customer has spent at Itikan in the past 2 years | Currency | 6 | $0.00-$9999.99 | No |

*Table 3: Data dictionary for Itikan’s marketing data.*

# Master Data

## What is Master Data?

Master data refers to core business objects used across various applications within an organisation. Examples of master data objects include the things that are logged into our transaction systems, measured and reported within reporting systems, and analysed in analytical systems. Key examples include customers, employees, vendors, suppliers, products, and locations (Loshin, 2010).

*Figure 5: Master data objects and what values are stored within them (Loshin, 2010).*

## Why is Master Data Important for a Company?

Proper management of master data prevents issues like data redundancy, inconsistencies, and misalignment of business processes. Implementing master data management creates a unified data foundation that supports informed decision-making, ultimately enhancing business operations (Roddewig, 2022).

## 4.3 Decision for Itikan E-Commerce Company

For Itikan, effective management of master data is vital to ensure that data objects contain accurate values, improving overall business operations. Below is a table defining Itikan’s master data objects, starting with the first row of the sample market data:

|  |  |
| --- | --- |
| **Market Data Object** | **Value Within Object** |
| Customer ID | 211 |
| Customer Name | Jenny Robinson |
| Recent Purchase | 15/5/2023 |
| Amount | $26.24 |
| Street Number | 3 |
| Street | Maple Avenue |
| State | New Zealand |
| Postcode | 624 |
| First Purchase | 14/8/2022 |
| Frequency 6 Months | 2 |
| Frequency 2 Years | 6 |
| Amount 6 Months | $52.48 |
| Amount 2 Years | $157.44 |

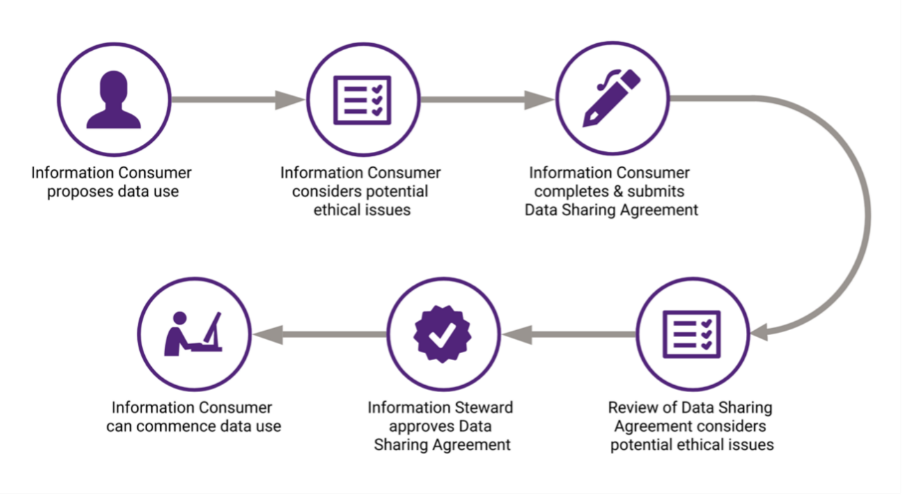
*Table 4: Example of master data management plan generated for Itikan*

# Data Ethics

## What are Data Ethics?

Data ethics refers to the principles a company establishes for gathering, protecting, and using customer data. According to (Floridi & Taddeo, 2016), it evaluates moral issues related to data generation, processing, and algorithms to support ethical conduct and values within the company.

## Why are Data Ethics Important?

Data ethics is crucial in a DMP for both stakeholders and customers. It ensures sensitive data is collected and shared legally and consensually. Adopting a data ethics framework enables companies to innovate while building trust that data is used responsibly (Australian Government, 2023).

*Figure 6: Steps in a data ethics framework (The University of Queensland, 2021)*

## 5.3 Decision for Itikan E-Commerce Company

Itikan needs a data ethics framework to ensure the safe, consensual, and legal handling of customer data. Below is the data ethics framework I’ve created for Itikan’s data:

|  |  |
| --- | --- |
| **Heading** | **Description** |
| Purpose and Transparency | * Provide a clear intent. * open communication with customers. |
| Data Minimisation | * Only collect necessary data. * Avoid sensitive data. |
| Consent | * Obtain explicit consent. |
| Data Security | * Encryption, access control, and cloud security. |
| User Privacy and Control | * Anonymisation. * User requests data deletion. |
| Data Sharing and Third Party Involvement | * Limited and purpose driven. |
| Compliance and Governance | * Regular internal and external inspections. * Legal compliance. |

*Table 5: Itikan’s data ethics framework guidelines.*

# Legal Compliance

## What is Legal Compliance

Legal compliance involves a company adhering to rules, policies, and procedures governing business practices. It requires understanding and implementing relevant legislation to ensure the company remains compliant throughout its processes (Donohue, 2023).

## Regulatory Frameworks, Data Protection Policies, and Audits

A regulatory compliance management framework includes modelling, monitoring, enforcement, and analysis of compliance with laws, contracts, and policies (Kharbili, 2012). Data protection policies are guidelines established to protect customer and business data. Audits evaluate whether a company’s data serves its intended purpose, ensuring unnecessary data is identified and eliminated per data minimisation principles.

## Decision for Itikan E-Commerce Company

Below is a regulatory compliance management framework regarding Itikan:

|  |  |  |
| --- | --- | --- |
| **Regulatory and Legal Requirement** | **Current State Within Itikan** | **Developed System and Processes** |
| Data Protection (e.g. General Data Protection Regulation GDPR). | Partially aligns with GDPR but lacks full compliance with data minimisation and consent. | Implement stricter data retention policies and user consent mechanisms. |
| Consumer Privacy (e.g. Consumer Data Right). | Policies in place but incomplete regarding consumer rights to access, delete, or opt-out of data sharing. | Develop a privacy portal for users to access, delete, and control their data sharing preferences. |
| Third-Party Data Sharing (e.g. California Privacy Rights Act CPRA). | No formalised process for reviewing third-party data-sharing agreements. | Implement formal review and approval workflows for third-party contracts to ensure data protection. |

*Table 6: Examples for regulatory compliance management framework for Itikan.*

# Storage and Backup

## Why is Storage and Backup Necessary?

Data storage is crucial for protecting vital information from loss due to hardware failure, natural disasters, or cyber-attacks. It also facilitates quick retrieval and sharing of information, enhancing collaboration and decision-making (OpsMatters, 2024). Meanwhile, data backups are equally as important, safeguarding against human errors, hardware failures, cyber-attacks, power outages, and natural disasters, which can save time and money. The best practices for data backups include:

* Understanding existing backup policies.
* Creating a backup policy if none exists.
* Backing up digital data and digitising physical documents.
* Automating backups.
* Choosing appropriate backup locations.
* Regularly checking backed-up files.
* Determining how long to keep backups.

(U.S. Geological Survey, 2023)

## Options for Storage and Backup

Storage Options (IBM, 2024):

* SSD and Flash Storage: Uses flash memory chips for fast data storage.
* Hybrid Storage: Combines the speed of flash storage with the capacity of hard disk drives.
* Cloud Storage: Stores data off-site, accessible via the internet or a private network.
* Hybrid Cloud Storage: Offers a mix of public, private, or both storage solutions.

Backup Options (Marget, 2021):

* Full Backup: Backs up all data to another location.
* Incremental Backup: Only backs up files changed since the last backup.
* Differential Backup: Backs up files changed since the last full backup.

## 7.3 Decision for Itikan E-Commerce Company

For Itikan, a hybrid cloud solution is the most optimal solution, combining public and private cloud features for secure and efficient data sharing. Recommended services include Amazon Web Services (AWS) Outposts, Oracle Cloud, or Microsoft Azure (The CTO Club, 2024). For backups, Itikan should start with a full backup to the hybrid cloud (e.g., Azure) to capture all initial data, followed by incremental backups for new or modified files. Full backups should occur monthly, on the 1st of each month, while incremental backups should be performed weekly, every Friday. This schedule balances thorough data protection with resource efficiency.

# Preservation

## What is Data Preservation?

Data preservation entails creating procedures to maintain data files over long periods using durable formats, local archiving, and repositories (University of Illinois, 2024). This ensures long-term accessibility. Effective preservation requires secure storage in hybrid cloud environments, across multiple locations (public and private), and in widely supported formats (e.g., CSV, XML, HTML). Data repositories, whether centralised or distributed, depend on the company's needs.

## How Long Should Data be Stored For?

There is no standard duration for data preservation, as it varies based on company requirements. For example, some data may need to be retained for five years, while other data might only require six months. Companies should develop custom data preservation plans tailored to the type and relevance of the data, ensuring efficient data management.

## Decision for Itikan E-Commerce Company

Below is a custom data preservation plan for Itikan’s sample market data, with tailored retention periods for each data type:

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Preservation Method** | **Frequency of Backup** | **Storage Location** |
| Customer ID | Encryption, Backup | Real-Time | Secure Database |
| Customer Name | Encryption, Backup | Real-Time | Secure Database |
| Recent Purchase | Archiving, Version Control | Daily | Cloud Storage, Local Servers |
| Amount | Encryption, Backup | Daily | Secure Database, Cloud |
| Street Number | Encryption, Backup | Weekly | Cloud, On-Premises Database |
| Street | Encryption, Backup | Weekly | Cloud, On-Premises Database |
| State | Encryption, Backup | Weekly | Cloud, On-Premises Database |
| Postcode | Encryption, Backup | Weekly | Cloud, On-Premises Database |
| First Purchase | Backup, Archiving | Weekly | Cloud Storage, Local Servers |
| Frequency 6 Months | Backup, Version Control | Weekly | Secure Database, Cloud |
| Frequency 2 Years | Backup, Version Control | Monthly | Secure Database, Cloud |
| Amount 6 Months | Encryption, Backup | Weekly | Secure Database, Cloud |
| Amount 2 Years | Encryption, Backup | Monthly | Secure Database, Cloud |

*Table 7: Itikan’s data preservation table.*

# Data Sharing

## What is Data Sharing?

Data sharing involves making data resources accessible to multiple applications, users, or external organisations, classified as internal or external entities. This process relies on technologies, legal frameworks, and cultural practices to ensure secure access while protecting data integrity. A key benefit is enhanced efficiency, fostering collaboration with vendors and partners (Amazon Web Services, 2024). According to (Waithira et al., 2019), credible and accurate data is essential for sharing while safeguarding the rights and confidentiality of participants.

## Internal Data Sharing

Internal data is generated within a company by departments like sales, finance, and marketing. This data is typically reliable for assessing internal performance but offers limited insight into the market or customers. While valuable, internal data can be narrow in scope, outdated, and lack diverse perspectives (Revelate, 2023).

## External Data Sharing

External data comes from sources outside the company, such as public records and market research. When effectively used, it enhances internal data, providing a comprehensive view of customers and the market. Companies leverage external data to identify trends, gain insights, and improve operations, ultimately leading to a better understanding of customers and new revenue opportunities (Revelate, 2023).

## Levels of Authority

Levels of authority refer to decision-making power and responsibilities within a company, defining who can make decisions and access data. In data management, this involves a plan specifying access based on roles within the company. This is crucial for managing internal data and sharing, distinguishing it from external data collection.

## Decision for Itikan E-Commerce Company

For Itikan, data sharing will categorise access types based on roles and responsibilities, as shown in the table below:

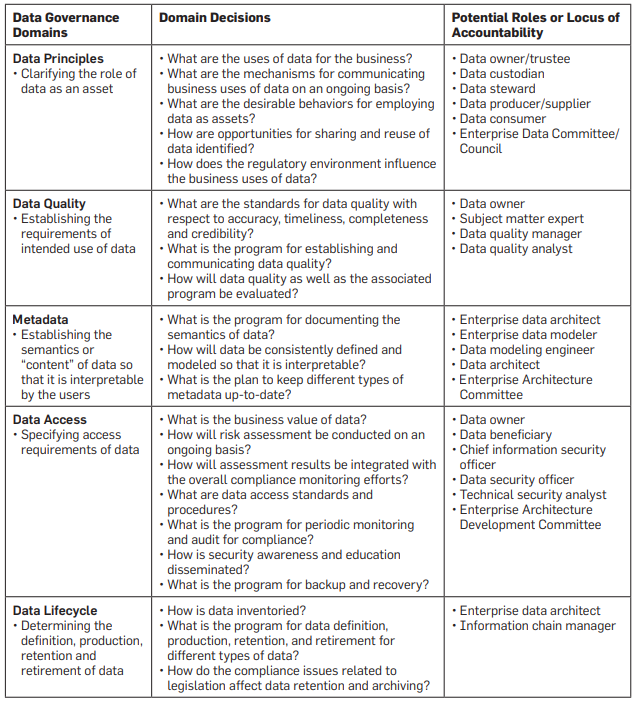
|  |  |
| --- | --- |
| **Data Access Type** | **Roles and Responsibilities** |
| Customer Information | * Sales and Marketing Team * Customer Service Representatives * Data Analysts |
| Purchase History and Frequency | * Sales, Marketing, Finance, and Data Analysts |
| Geographical Information | * Marketing, Logistics, and Customer Service Teams |
| Frequency and Amount of Purchases Over Time | * Data Analysts * Finance and Sales Teams |

*Table 8: Itikan’s data sharing roles and responsibilities.*

Data sharing will employ role-based access control (RBAC), granting access to specific data sets based on job roles. For instance, customer service representatives can access customer information, while finance analysts cannot. To secure data sharing, measures like encryption and access logs will be implemented, ensuring only authorised personnel access sensitive data. This approach enhances data sharing security and ensures compliance with data privacy regulations.

# Data Governance

## What is Data Governance?

Data governance is essential for establishing decision-making authority and accountability regarding a company’s data assets (Khatri & Brown, 2010). Their framework identifies five interconnected domains: data principles, data quality, metadata, data access, and data lifecycle. These domains interact, starting with data principles that set usage boundaries, which inform standards for data quality, interpretation (metadata), and user access (data access). Additionally, data lifecycle decisions, like creation, retention, and retirement, are influenced by these principles (Khatri & Brown, 2010).

*Figure 7: Data decision domains commonly used within data governance (Khatri & Brown, 2010).*

## Data Policies

Policies are vital in a data governance plan, defining rules for data usage, access permissions, and allowed operations. They clarify procedures, responsibilities, and management processes, ensuring control and accountability in data practices (Atlan, 2024). Six key policies from (Atlan, 2024) include:

* Data Quality: Ensures data is suitable for its purpose by assessing validity, completeness, timeliness, accuracy, and integrity.
* Data Privacy: Establishes accountability and transparency, minimising privacy risks while complying with regulations.
* Data Security: Manages access to valuable information, reducing security risks.
* Data Lifecycle: Provides guidelines for data retention, archival, and deletion.
* Data Ethics: Aligns data management and collaboration with company values.
* Data Definitions and Models: Standardises important concepts, terms, master data, and key performance indicators (KPIs).

## Data Standards

Data governance standards help define data definitions, taxonomies, master data, and enforce technical standards within an organisation (Panian, 2010). These standards serve as guidelines for managing the quality, availability, usability, integrity, and security of data assets (Atlan, 2023). From (Atlan, 2023), they further identify eight key standards commonly used in data governance: data quality, data integration, data security and privacy, data accessibility, data lifecycle management, data compliance, data ownership and stewardship, and metadata management.

## Decision for Itikan E-Commerce Company

I have developed a comprehensive table outlining the roles and responsibilities of individuals at Itikan, using key information from (Data.NSW, 2024). This table summarises various roles within Itikan, detailing specific actions and responsibilities related to data governance. It provides a clear overview of how data governance is structured, highlighting accountability and duties in managing the company’s data effectively, securely, and in compliance with regulations.

|  |  |  |
| --- | --- | --- |
| **Role** | **Description** | **Key Responsibilities** |
| Accountable Executive | Senior executive with ultimate accountability for the data, often delegated to the Data Sponsor/Data Owner. | |  | | --- | |  |  |  | | --- | |  |   - Approve policies, protocols, and guidelines.  - Ensure legal, regulatory, and policy compliance.  - Approve significant changes to data collection/process/system.  - Approve funding for data management projects.  - Monitor data governance performance.  - Ensure data assets are identified and documented.  - Delegate responsibilities to Responsible Executives. |
| Responsible Executive | Director-level role with delegated responsibility for a specific data asset, typically the Data Custodian. | - Enforce rules on behalf of Accountable Executive. - Document data assets in a data catalogue. - Identify information security classification. - Determine conditions for data use, storage, and sharing. - Approve data change, sharing, and release requests. - Set standards for the data asset. - Nominate and oversee the Operational Data Manager. - Develop strategic plans for data asset management. |
| Operational Data Manager | Business manager or subject matter expert responsible for operational data management, sometimes referred to as the Data Steward. | - Day-to-day management of data asset. - Ensure data sharing agreements/license compliance. - Manage data according to legislation, policies, and standards. - Maintain metadata, data dictionary, and business rules.  - Advise Responsible Executive on asset management. - Guide Data Users on proper data use and interpretation. - Provide feedback to Data Creators on data quality issues. |
| Data Creator | Employee, contractor, or consultant who captures or creates data for the agency, also known as the 'Supplier'. | - Collect data according to agreed standards. - Provide accurate and detailed metadata. - Maintain ongoing data processes. - Ensure data security. - Comply with legislation, policies, and terms of collection. - Ensure consent is obtained where applicable. |
| Data User | Anyone (public or government) who uses data. | - Acknowledge data source and comply with copyright or licensing. - Ensure data is fit for intended purpose. - Report data errors or omissions. - Comply with access license/agreement terms. - Maintain security and privacy of data. - Report breaches to Operational Data Manager or Responsible Executive. - Obtain approval for data release from Accountable Executive. |

*Table 9: Roles and responsibilities managed in data governance.*

# Data Security

## What is Data Security?

Data security involves methods used to protect a company’s information, including access controls, data encryption, and multi-factor authentication. Formally, it is the practice of safeguarding digital information from unauthorised access, corruption, or theft throughout its lifecycle (IBM, 2021). This includes physical protection of hardware, administrative controls, and the logical security of software applications, supported by organisational policies and procedures to ensure secure data handling (IBM, 2021).

## Data Security Strategies

To defend against cyber threats, companies should create a tailored data security plan. Effective data security not only protects information but also builds customer trust through transparency, integrity, and secure data processing (Cloudian, n.d.). Key strategies for a robust data security plan include:

* Data lifecycle management
* Data risk management
* Data backup and recovery
* Data access management
* Data storage management
* Data breach protection
* Confidentiality, integrity, and availability
* Data protection policies and procedures
* Standards and regulatory compliance
* Monitoring and auditing

(Cloudian, n.d.)

## Decision for Itikan E-Commerce Company

For Itikan, developing a customised data security plan is essential to safeguard its data in today’s digital environment. By implementing the ten outlined strategies, Itikan can effectively protect its sample market data from cybersecurity threats. Among the provided data, the most sensitive information includes Customer ID, Name, Street Number, Street, State, and Postcode. These fields are highly confidential, containing critical personal details. Conversely, other data fields such as Recent Purchase and Frequency are transactional and do not pose the same sensitivity risk as personal customer data.

# Data Disposal

## Why is Data Disposal Important?

Data disposal refers to the transfer of data ownership or destruction of data on electronic media, making it unreadable and inaccessible for unauthorised use. Once deleted, data cannot be accessed by the operating system or software tools (Bigelow, 2024). In data lifecycle management, proper disposal is crucial for security, especially for business sensitive data that contains personal information or is subject to regulations. Disposing of data when its lifecycle ends is vital for maintaining data security (Bigelow, 2024).

## How is Data Organised for Disposal?

Organising data for disposal requires adherence to minimum retention periods set by an approved records authority. The process includes arranging for secure destruction or transferring data ownership to another entity, like governmental agencies or national archives. Transferred data must meet the receiving agency's requirements and be formatted for long-term preservation (National Archives of Australia, n.d.).

## Decision for Itikan E-Commerce Company

For Itikan’s sample market data, a disposal table outlines retention periods and sensitivity for each data type. This helps Itikan understand what data should be deleted or transferred:

|  |  |  |
| --- | --- | --- |
| **Data Type** | **Retention Period** | **Sensitivity** |
| Personal customer data | 5 years after last interaction | High |
| Financial data | 7 years | Medium |
| Employee data | 7 years post-termination | High |
| Marketing data | 2 years | Low |

*Table 10: Data disposal table.*

Itikan will dispose of unwanted data through inventory tools to identify eligible data for disposal. Approved data destruction software will permanently delete digital data in compliance with standards. Physical data disposal will involve identifying documents for destruction and employing certified methods to securely destroy sensitive materials.

# Data Integration

## How Does Data Integrate?

Data integration involves consolidating data from multiple sources to create a complete, accurate, and up to date dataset for business intelligence and analysis. The primary method for integration is the ETL process, which standardises data for storage in a repository. Other integration methods include ELT, data streaming, application integration, and data virtualisation (Qlik, 2024).

## Why is Data Integration Necessary?

Data integration is essential for enabling analysis and business intelligence, as it standardises data formats. It provides a holistic view of data, enhances decision-making, streamlines processes, improves data accuracy, fosters collaboration, ensures regulatory compliance, and offers better customer insights and scalability. Without integration, companies miss out on these benefits, potentially sacrificing profitability.

## Decision for Itikan E-Commerce Company

Itikan currently uses the ETL process for data integration, as seen in its sample market dataset. To enhance capabilities, the company should consider collecting additional external data while maintaining the ETL process. For future investments, adopting a database system like Oracle or MySQL could streamline operations and reduce reliance on ETL. Though these databases may involve initial costs and technical challenges, they offer significant long-term benefits in data storage, security, and organisation.

# Monitoring and Reviewing of Data

## Why is it Important to Monitor and Review data

Monitoring and reviewing data is crucial for ensuring its accuracy and relevance. Regular assessments lead to higher data quality, resulting in better data-driven decisions, enhanced competitive advantage, improved understanding of customer needs, and increased operational efficiency (Chu, 2024).

## Decision for Itikan E-Commerce Company

At Itikan, data monitoring can be streamlined by establishing a weekly review schedule. A senior team member should be tasked with assessing all critical data on a designated day each week, ensuring it remains current and secure, thus optimising future data usage.

# Conclusion

In conclusion, the DMP for Itikan covers the essential areas of data collection, metadata, ethics, legal compliance, storage and backup, preservation, data sharing, and governance. Additionally, it includes sections on master data, data security, disposal, integration, and monitoring. Together, these components enhance Itikan’s data management strategy, supporting the company’s growth and improving its internal and external business processes as it expands.

# References

* Alexsoft. (2022). Metadata management: Process, tools, use cases, and best practices. Alexsoft Software R&D Engineering. https://www.altexsoft.com/blog/metadata-management/
* Amazon Web Services. (2024). What is Data Sharing? - Data Sharing Explained - AWS. Amazon Web Services, Inc. https://aws.amazon.com/what-is/data-sharing/
* Atlan. (2023). 8 data governance standards your data assets need now. Atlan.com. https://atlan.com/data-governance-standards/
* Atlan. (2024). Data governance policy: Examples, templates & how to write one. Atlan. https://atlan.com/data-governance-policy/
* Australia Research Data Commons. (2022, May 14). Data management plans. https://Ardc.edu.au/. https://ardc.edu.au/resource/data-management-plans/
* Australian Bureau of Statistics. (2014, September 30). Foundation for a national data collection and reporting framework for family, domestic and sexual violence, 2014 | australian bureau of statistics. https://www.abs.gov.au/statistics/people/crime-and-justice/foundation-national-data-collection-and-reporting-framework-family-domestic-and-sexual-violence/2014
* Australian Government. (2023). Data Ethics Framework (New). Data and Digital. https://www.dataanddigital.gov.au/plan/roadmap/trusted-and-secure/data-ethics-framework
* Baker, N. (2023, July 30). What is the main purpose of data collection? - savanta. Savanta.com. https://savanta.com/knowledge-centre/view/what-is-the-main-purpose-of-data-collection/
* Bigelow, S. (2024). What is data destruction? - definition from whatis.com. Tech Target. https://www.techtarget.com/searchstorage/definition/data-destruction
* CESSDA. (2017). Benefits of data management - data management expert guide. Dmeg.cessda.eu. https://dmeg.cessda.eu/Data-Management-Expert-Guide/1.-Plan/Benefits-of-data-management
* Chu, D. (2024). What is data monitoring and why does it matter? | secoda. Www.secoda.co. https://www.secoda.co/blog/what-is-data-monitoring-and-why-is-it-important
* Cloudian. (n.d.). Data protection strategy: 10 components of an effective strategy. Cloudian. https://cloudian.com/guides/data-protection/data-protection-strategy-10-components-of-an-effective-strategy/
* Cunningham, A. (n.d.). ISO 23081 Metadata for records. Committee.iso.org. https://committee.iso.org/sites/tc46sc11/home/projects/published/iso-23081-metadata-for-records.html
* Data.NSW. (2024). Module 6: Assigning roles and responsibilities | data.nsw. Data.nsw.gov.au. https://data.nsw.gov.au/data-governance-toolkit-0/module-6-assigning-roles-and-responsibilities
* Donohue, J. (2023, January 10). What is legal compliance? Diligent.com. https://www.diligent.com/en-au/resources/blog/steps-evaluating-legal-compliance
* Floridi, L., & Taddeo, M. (2016). What Is Data ethics? Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2083). https://doi.org/10.1098/rsta.2016.0360
* IBM. (n.d.). What is data management plan? | IBM. Www.ibm.com. Retrieved October 12, 2024, from https://www.ibm.com/topics/data-management-plan
* IBM. (2021). What is data security? Definition, solutions and how to secure data. IBM. https://www.ibm.com/topics/data-security
* IBM. (2024). What is data storage? Www.ibm.com. https://www.ibm.com/topics/data-storage
* Informatica. (2024). What is ETL? Informatica. https://www.informatica.com/tw/resources/articles/what-is-etl.html
* Khatri, V., & Brown, C. V. (2010). Designing data governance. Communications of the ACM, 53(1), 148. https://doi.org/10.1145/1629175.1629210
* Kharbili, M. (2012). Business Process Regulatory Compliance Management Solution Frameworks: A Comparative Evaluation. https://crpit.scem.westernsydney.edu.au/confpapers/CRPITV130Kharbili.pdf
* Lang, A. (2024). Data management practices hierarchy. Canvas. https://learn.usc.edu.au/courses/27682/pages/module-9-topic-3-data-strategy?module\_item\_id=596449
* Loshin, D. (2010). *Master Data Management*. Google Books. https://books.google.com.au/books
* Marget, A. (2021, October 19). Types of Backups: Full, Incremental and Differential Backup. Unitrends. https://www.unitrends.com/blog/types-of-backup-full-incremental-differential
* Mayernik, M. (2015). *Metadata (IEKO)*. Isko.org. https://www.isko.org/cyclo/metadata#3.1
* Michener, W. K. (2015). Ten simple rules for creating a good data management plan. PLOS Computational Biology, 11(10), e1004525. https://doi.org/10.1371/journal.pcbi.1004525
* National Archives of Australia. (2015). *Australian government recordkeeping metadata standard (AGRkMS)*. https://www.naa.gov.au/sites/default/files/2019-09/AGRkMS-Version-2.2-June-2015\_tcm16-93990\_1.pdf
* National Archives of Australia. (n.d.). Retaining, managing and disposing of data and datasets | naa.gov.au. Www.naa.gov.au. Retrieved October 21, 2024, from https://www.naa.gov.au/information-management/disposing-
* OpsMatters. (2024, February 12). Data Storage: What Is It and Why Is It Important? | OpsMatters. Opsmatters.com. https://opsmatters.com/posts/data-storage-what-it-and-why-it-important
* Panian, Z. (2010). Some practical experiences in data governance . World Academy of Science, Engineering and Technology, 62(1), 8.
* Qlik. (2024). What is data integration? Definition, examples & use cases. Qlik. https://www.qlik.com/us/data-integration
* Revelate. (2023, October 17). External and internal data differences in the data marketplace. Revelate. https://revelate.co/blog/external-and-internal-data-differences-in-the-data-marketplace/
* Roddewig, S. (2022, August 23). Master Data: What Is It & Why Does It Matter for Businesses? Blog.hubspot.com. https://blog.hubspot.com/marketing/master-data
* Sen, A. (2004). Metadata management: Past, present and future. *Decision Support Systems*, *37*(1), 151–173. https://doi.org/10.1016/s0167-9236(02)00208-7
* Singh, R. K., & Madalli, D. P. (2023). DMPFrame: A conceptual metadata framework for data management plans. *Journal of Library Metadata*, *23*(3-4), 121–160. https://doi.org/10.1080/19386389.2023.2268474
* The CTO Club, T. C. (2024). 22 Best Hybrid Cloud Solutions Reviewed For 2024. https://thectoclub.com/tools/best-hybrid-cloud-solutions/
* The University of Queensland. (2021, April 14). Enterprise Data Ethics Framework. Data.uq.edu.au. https://data.uq.edu.au/enterprise-data-ethics-framework
* Thoegersen, J. (2015). Examination of federal data management plan guidelines. Journal of EScience Librarianship, 4(1). https://doi.org/10.7191/jeslib.2015.1072
* University of Illinois. (2022). LibGuides: Introduction to Data Management for Undergraduate Students: Data Preservation. Guides.library.illinois.edu. https://guides.library.illinois.edu/introdata/preservation
* U.S. Geological Survey (2023) USGS Data Management Website: U.S. Geological Survey, https://doi.org/10.5066/F7MW2G15
* Waithira, N., Mutinda, B., & Cheah, P. Y. (2019). Data management and sharing policy: The first step towards promoting data sharing. BMC Medicine, 17(1). https://doi.org/10.1186/s12916-019-1315-8

# Appendix

## Appendix 1 (Sample Marketing Data)

